Author Response to Reviewer #1

The authors thank Reviewer #1 for their thorough review. We have addressed your comments. Please see the response to each point below in red. Figures in this reply are ordered with capital letters to distinguish them from the figures in the manuscript.

Changes to the paper not discussed in the Authors' Comments:

- MLCs above 0°C are treated differently in the observational algorithm and were not constrained by a 150m gap; this has now been rectified, and values have thus changed in Fig.7. Also, values for the RS product have changed due to an update in the observational algorithm. Please see Fig. A and B below.
- There were some inconsistencies in the development of the two algorithms; this has now been rectified. All model data has been updated with a 150m gap threshold. Overall, small changes are induced (MLC occurrence for 1E-9 kg/kg goes from 77% to 76%).
- Updated acknowledgements to follow the Supercomputer HoreKa suggested structure
- Updated Fig. A1 with [ ] brackets instead of ( ) for the units
- Wrong units in Fig. B1b

Specific Comments

- My major concern with this study concerns the intercomparison of MLC occurrence and cloud properties between the model and the observations. The observationally based occurrence is derived from radiosondes (RS and further supplemented by cloud radar observations (RS+Radar), whereas MLC detection in the model is based on cloud mass.

  The algorithm also asserts supersaturation with respect to ice or water (Line 161).

  The authors clearly demonstrate the sensitivity of MLC occurrence to the chosen cloud mass threshold. Ultimately, they select a threshold of 10-9 kg kg-1for comparison with RS detections. However, the rationale for using RS-only detection rather than RS+Radar, which reduces the detection of spurious cloud layers, is unclear. Furthermore, it is not evident why the 10-9 kg kg-1 threshold was chosen, given that a threshold of 10-8 kg kg-1 appears to more closely match RS detections.

  Yes, this choice may seem arbitrary. We chose 1E-9 kg/kg to follow the model's limit for radiation to interact with the clouds. However, we have now changed this following comments from Reviewer #2 and will consider the cloud mass threshold of 1E-5 kg/kg mainly, but also showing medians for cloud height, thickness, and gaps for the other thresholds to remove the arbitrary part of the comparison.

We further add the RS+Radar cloud heights and thicknesses for the model comparison instead of the RS only, and include all cloud mass thresholds (CMTs) to remove this absolute comparison to a certain CMT.

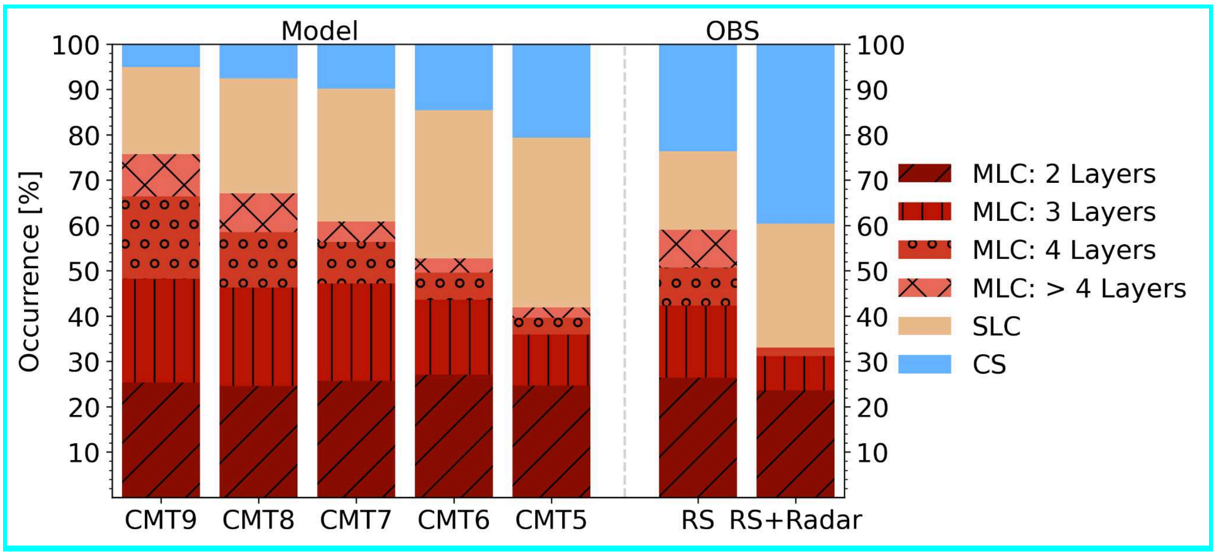

Fig A (Fig.7) with new colours, added CMT5 and updates to the observational product.

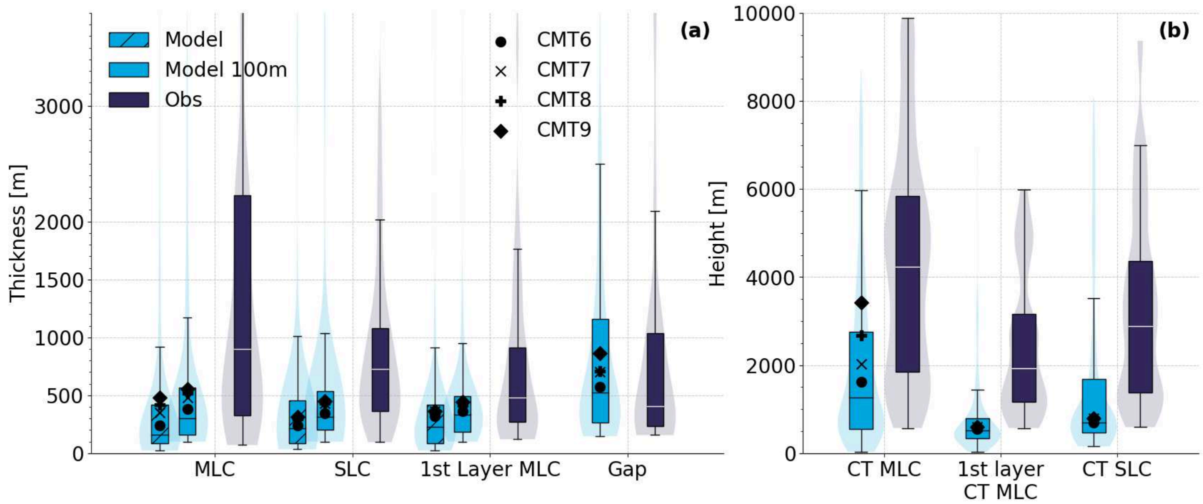

Fig. B (Fig.8) Cloud thickness and cloud gap thresholds are more consistently treated in the RS+Radar product. CT heights are shifted up due to the removal of cloud layers below the lowest radar gate (see Achtert et al. (2025) for details). Outlines of the distributions are added on request from Reviewer #2. A filtered model data category has been added, where clouds with a thickness of less than 100m are excluded.

Or is it because using a low threshold would ultimately lead to RS-like definition as only the saturation criterion is considered? In such a case, it would be important to have a consistent definition of saturation (see also bullet point 3).

Please see the answer below regarding the saturation definition.

Similar concerns apply to the comparison of seeding occurrence (Tab. 1).

We have now also added the seeding occurrence for all cloud mass thresholds to be less deterministic in terms of choosing a “correct” cloud mass threshold

One potential way to avoid the need to arbitrarily select a cloud mass threshold would be to employ a radar forward operator to generate radar reflectivity, enabling a more consistent comparison with the Achtert et al. (2025t detection algorithm.

Yes, this would have been a good idea to implement. Unfortunately, we do not have all the output parameters required to employ a radar forward operator at this stage, and new simulations cannot be performed due to the computational constraints. We hope that the further comparison with threshold 1E-5 kg/kg and adding the uncertainty (in terms of all mass thresholds) in the following analysis addresses the reviewer’s concerns.

- **P7, L160**: It took some time to realize that two distinct thresholds are used in this study: a cloud mass threshold and a seeding mass threshold, which share the same numerical values. It would be helpful to clearly distinguish these thresholds throughout the manuscript, for example, by using separate mathematical symbols, to avoid confusion.

  Yes, this is a fair point. We have updated the manuscript to refer to a cloud mass threshold (CMT) and a seeding mass threshold (SMT) throughout the text.

  Additionally, it is unclear why the cloud mass threshold changes between sections (10-6 kg kg-1 in Section 5.1 vs. 10 -9 kg kg-1 for MLC detection). While exploring the sensitivity of MLC properties to this threshold is valuable, once a threshold is chosen, it would be advisable to report all other microphysical properties (cloud droplet number concentration, …) using the same threshold, unless there is a compelling reason not to do so.

  We used our previous greatest cloud mass threshold to better ascertain whether any changes were seen within the clouds with the new parameterisations. We generally agree on consistency; however, here we are evaluating whether there is a general response to the change in parameterisations rather than for a cloud mass threshold that is tuned to observations. We can, however, give a range of values. Thus, we evaluate CMT5 and CMT9 for the two parameterisations and plot these in Fig. C. We find a 2%-16% difference in cloud ice at warm temperatures with CMT5 and CMT9, respectively. We update the section accordingly.

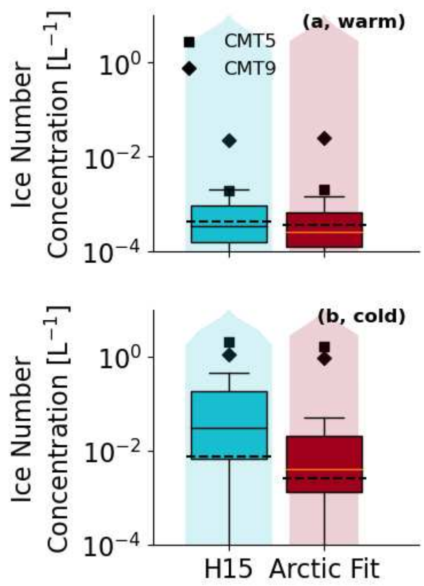

Fig. C (Fig. 3) The boxplots show the distribution of CMT5 and mean (diamonds) and median (dashed line) for CMT9. Violin plots (on request from Reviewer #2) show the distribution of CMT5.

P7, L161–162: How do the authors decide with respect to which phase (liquid or ice) saturation is determined? Is this approach consistent with the method used to calculate saturation for cloud cover in ICON or with Achtert et al. (2025)?

For both algorithms, both saturations are taken into account. The clause is implemented with an OR statement such that when saturation for either liquid or ice is reached, the layer is flagged as a cloud layer. In the model, this is then combined with a cloud mass threshold (LWC+IWC) such that we ensure cloud mass is present. This allows for both fully liquid, fully ice, and mixed-phase clouds to be flagged. In the observational algorithm, the presence of clouds is validated using radar. The model approach is similar to Achtert et al. (2025), but the measured variables come with uncertainty that we do not take into account here. We have not aligned the algorithm with the cloud cover scheme in ICON.

In this context, it would also be helpful to clarify whether the model employs a fractional or grid-scale cloud cover scheme, as this is not explicitly indicated in the model description.

For the radiation calculation, we use a diagnostic cloud cover, a type of fractional cloud cover scheme. In the microphysics scheme, a simple grid-scale approach is used (0 or 1).

- **P9, L219–241:** The reported values for mean cloud droplet number concentration (Ndt appear unusually high for the Arctic, even exceeding the number of observed cloud condensation nuclei that could potentially be activated (see Fig. 2). Could this be due to the mean being influenced by outliers, as seems to be the case for mean ice crystal number concentrations (see Fig. 3)?

Yes, thank you for catching this. There was a bug in the calculation for the units. We update the section with a table showcasing the mean and median number concentrations for cloud liquid and ice. The largest mean is 40 cm^-3, which is more in line with what the parameterisation allows for.

A similar concern applies to the reported cloud ice masses and number concentrations. I would suggest that reporting median values may provide a more robust representation of these quantities. Do the results differ when evaluating medians instead of means?

The medians are indicated in Fig. 3 by the horizontal lines, and as stated in Line 231 these medians remain larger for H15 than for the Arctic fit. We only find larger values in the mean values due to these outliers, as you also mention. Thus, as we try to state, the impact of the parameterisation is limited.

We also update this section with the CMT5-CMT9 threshold and the respective new values in the table mentioned above. The new Figure 3 is shown above in Fig. C.

- **P10, Fig. 3:** In the caption, you state that values outside the interquartile range (IQR) are excluded, yet these values still appear to be included when calculating the means shown in the figure and subsequently reported in the manuscript. This reinforces the earlier concern: using medians would make the reported statistics less sensitive to outliers and might remove the need to filter extreme values in the first place.

  A slight miswording, the outliers are only visually excluded to aid the interpretation of the figure. We add to each figure with boxplots:

  "Values larger (or smaller) than these are outliers and are not shown (only in the figure) to simplify the visual interpretation."

- **P13, Fig. 5:** As stated by the authors, liquid water content is not given in CloudNet if liquid-containing clouds have liquid-phase precipitation. I wonder how the median liquid water path has been derived for the model and for Shupe–Turner. Were time steps with liquid-phase precipitation excluded from the comparison? If not, this may lead to a definition-inconsistent intercomparison, as the rainwater path is included in the model output and in Shupe–Turner, but not in CloudNet.

  The ShupeTurner algorithm employs a similar method as they're bound by the same observations. We add to methods:

  "LWC is not available during liquid-phase precipitation in both of these retrievals."

We also update the LWP statement as we are currently using the integrated LWC and not the measured LWP.

Line 187: **"Liquid water path (LWP) and ice water path (IWP) are calculated as the column-integrated LWC and IWC, respectively. The uncertainty in LWC is 15% to 25%. "**

The comparison to observations is always difficult. We included the rainwater path in the modelled LWP, as this is rain within the clouds and not precipitation reaching the surface. For simplicity, we may remove it. The median is marginally reduced by about 6% and the factor differences remain the same.

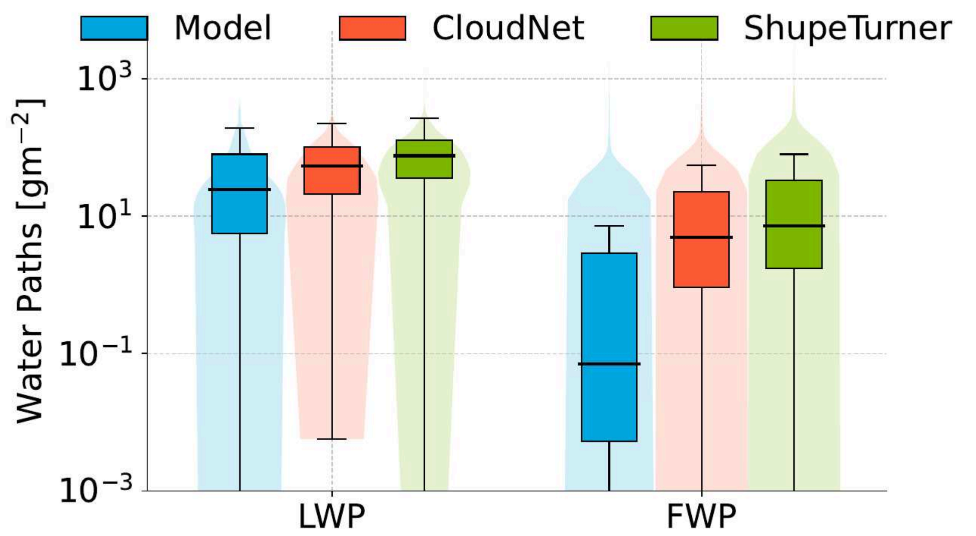

Fig D, (Fig. 5), where LWP now only contains liquid water and no in-cloud rain.

P22, L463–464: Could you provide more information about the physical pathway responsible for the increase in geometrical cloud thickness?
To be clear, there is no "increase" in the thickness; rather, we find that MLCs are thicker than SLCs. The 'strengthening' we are hypothesising refers to this finding. For simplicity and to refrain from making any hypotheses in the manuscript, this sentence is removed.

Minor Remarks

- **P2, L39–41:** While seeding can indeed initiate glaciation, neither riming nor secondary ice production can initiate it, since both processes require pre-existing cloud ice. These processes should therefore be described as enhancing glaciation rather than initiating it. Similarly, the current phrasing suggests that the Wegener–Bergeron–Findeisen (WBF) process initiates glaciation, whereas it primarily enhances glaciation once cloud ice is present. Consider rewording this section.

Perhaps the word "initiate" is a bit misleading here. We have changed this to:

“Glaciation, the transition from a mixed-phase state to fully ice, may be enhanced by the seeding of frozen precipitation (ice crystals, snow, or graupel) together with riming and secondary ice production (SIP), through the Wegener-Bergeron-Findeisen (WBF) mechanism”

- **P2, L50–51:** Downwelling longwave radiation will only influence the lower cloud layer and not “each other.”

  Changed to **“Radiatively, overlaying cloud layers in MLC systems influence lower clouds through an increase in downwelling longwave radiation.”**

- **P3, L72–73:** “ICON Global analysis”: Are you referring to the analysis step (0th timestep) of the global forecast here? If so, I wonder whether this analysis is produced every 3 hours, as you further down state that you employ boundary conditions with 3-hourly updates.

  Yes, we initialise the model at 00UTC (Line 87) from the analysis. We rewrite Line 74 to make the product clearer:

  “The ICON Global analysis is a combination of forecast and data assimilation. Every 3 hours, a new data-assimilation cycle is initiated using the global observing network and local data assimilation from the radiosoundings during MOSAiC. Thus, we maintain a close agreement to observations at initialisation and boundary conditions supply changes along the domain edge with 3-hourly updates. ”

- **P3, L74–75:** Here, one might understand that radiosondes are used as the only observations during the data assimilation. I assume you refer to the fact that the radiosonde observations during MOSAiC are assimilated, in addition to the standard global observations. Furthermore, are observations really nudged (which I consider some kind of Newtonian relaxation) or simply used during the data assimilation step?

  No, they’re simply used in the data assimilation. Thank you for catching this misuse of the word. The rewrite is listed above, P3, L72-73.

- **P7, L165–166:** No need to repeat the conditions, as you are already referring to them in the first part of the sentence.

  Ok, removed.

- **P8, L188–189:** “A standard deviation of the mean.” Do you mean that the standard deviation is the same magnitude as the mean?
  We excuse the ambiguity, this has now been removed to also better reflect the fact that we are currently using the integrated LWC and not the measured LWP.

Line 187: **“Liquid water path (LWP) and ice water path (IWP) are calculated as the column-integrated LWC and IWC, respectively. The uncertainty in LWC is 15% to 25%. “**

- **P20, L413–414:** Or because you are in an updraft-limited regime. On this end, I assume that grid-scale vertical velocity is used for aerosol activation, which might be too low at kilometer-scale resolution for Arctic clouds, which might be turbulence-driven.

  That is a very good point. We add: **“It may also be due to too low vertical velocities, limiting the cloud droplet activation.”**

- **P21, L448:** “... during the aircraft campaign PS106 ...” Isn’t PS106 a ship cruise?

  Yep, you’re very correct. Thanks for catching this. We update:

  “At lower latitudes, close to Svalbard, during the PS106 campaign, an occurrence of 36% of MLCs was reported…”